DOCUMENT RESUME

ED 288 909 TM 870 716

AUTHOR Tal, Joseph S.
TITLE Numerical Facilities: A Review of the Literature. Technical Report 1985-3.
INSTITUTION Johnson O'Connor Research Foundation, Chicago, IL. Human Engineering Lab.
PUB DATE Dec 85
NOTE 20p.
PUB TYPE Information Analyses (070) -- Reports - Research/Technical (143)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Arithmetic; Cognitive Ability; Cognitive Measurement; *Cognitive Tests; Factor Structure; *Intelligence; Literature Reviews; *Mathematics Skills; *Mathematics Tests; *Nonverbal Ability
IDENTIFIERS Army Air Forces Classification Batteries; Army Classification Battery; Thurstone Primary Mental Abilities Schema

ABSTRACT

This review of the relevant literature in the area of numerical facility attempts to clarify the construct of numerical facility and provide guidance for items tapping this ability. The review is presented in five parts. The first section introduces two approaches that can be used to investigate numerical facility, including factor analysis. Studies that have attempted to isolate a numerical factor are presented in the second section. The third section deals with mathematical ability, a construct typically viewed as related to numerical facility. Several studies relating numerical facility and "automatic processes" are presented in the fourth section. The fifth section supports Thurstone's (1938) conclusion that the ability to manipulate numbers is a unique ability and that there is some evidence that this is genetic in nature. Although most of the studies measuring this ability used problems involving simple arithmetic, this may not be the best way to measure numerical facility. It is suggested that tasks involving the manipulation of symbols other than numerals might be preferred. A recommendation is made that the Johnson O'Connor Research Foundation use three tests to research this construct: (1) simple arithmetic; (2) simple arithmetic involving some arithmetic reasoning; and (3) nonnumerical manipulation. (KSA)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED288909

NUMERICAL FACILITIES:

A REVIEW OF THE LITERATURE

by
Joseph S. Tal

JOHNSON O'CONNOR RESEARCH FOUNDATION
HUMAN ENGINEERING LABORATORY

Technical Report 1985-3
December, 1985

BEST COPY AVAILABLE

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☒ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
R. F. Kyle

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

TM 870 716

Numerical Facilities: A Review of the Literature

by
Joseph S. Tal

JOHNSON O'CONNOR RESEARCH FOUNDATION

HUMAN ENGINEERING LABORATORY

Technical Report 1985-3

December. 1985

COPYRIGHT © 1986 BY JOHNSON O'CONNOR RESEARCH FOUNDATION, INCORPORATED

Numerical Facility: A Review of the Literature

by Joseph S. Tal

Abstract: The purpose of this report, which reviews the relevant literature in the area of numerical facility, is to clarify the construct of numerical facility and to provide guidance for developing items tapping this ability.

It has been concluded that there indeed exists an ability that is uniquely associated with the manipulation of numbers. Most of the studies that have measured this ability have used problems involving simple arithmetic. However, this may not be the only way, or even the best way, to measure numerical facility. It is suggested that tasks involving the manipulation of symbols other than numerals may present a better way of measuring this ability. A recommendation is made that the Johnson O'Connor Research Foundation use three tests to research this construct.

Introduction

This paper presents a review of the literature in the area of numerical facility. The review is presented in five parts. The first section introduces two possible approaches that can be used to investigate numerical facility. Studies that have attempted to isolate a numerical factor are presented in the second section. The third section deals with mathematical ability, a construct that is typically viewed as related to numerical facility. Several studies relating numerical facility and "automatic processes" are presented in the fourth section. The conclusions that appear warranted at this time are presented in the fifth and final section.

I. Approaches to testing numerical facility

On a general level, there are at least two approaches to constructing items tapping any mental ability. The first involves an initial conceptualizing of the nature of the ability and the construction of items consistent with that formulation. Such an approach requires a fairly precise formulation of the construct under study. Tests that result from this approach are expected to contain a homogeneous set of items. In the second approach, a fairly heterogeneous set of items is developed and administered. Data reduction techniques (e.g., factor analysis, principal component analysis, etc.) are then applied in the hope that item intercorrelations will reveal useful information about the construct.

The first approach is confirmatory in nature while the second is exploratory. Exploratory studies do not require as precise a theoretical formulation as confirmatory studies. However, some theory is needed in order to dictate which tests are to be used in the study. At this juncture it is useful to provide a brief overview of factor analysis since it is this technique that has been used most often in investigating numerical facility.

Factor Analysis

Factor analysis is one of a family of techniques that attempt to "simplify the description of data by reducing the number of necessary variables or dimensions" (Anastasi, 1976; p. 362). For example, an experimenter may be interested in the possibility that there are different types of reasoning. One way to go about researching this is to administer a relatively large number of tests that involve reasoning and to compute the correlations among these tests. If the intercorrelations among these tests are similar in magnitude, the experimenter will be led to believe that one reasoning dimension is responsible for the relationships among these tests. If, for example, a number of tests correlate with one another but not with other tests, the experimenter will

conclude that the intercorrelated tests share a common element that is not shared with the other tests. Factor analysis is a statistical procedure designed to discover patterns such as these within a correlation matrix consisting of pairwise correlations between tests or between items within a single test.

The basic assumption of factor analysis is that there exist underlying factors, or dimensions, that can explain the observed intercorrelations among variables. These factors can be revealed by the pattern of correlations among variables. The correlation between a variable and a factor is called a factor loading. The greater the factor loading, the greater the amount of variance in the variable that is explained by the factor.

A discussion of the statistical procedures that yield factors is beyond the scope of this paper. However, it is important to note that subjective decisions are involved at almost every step in a factor analytic study. This includes the initial stages involving the selection of tests and individuals, as well as the later stages of statistically extracting factors from the data and the naming of these factors. Consequently, results of factor analytic studies may sometimes be specific to the particular study at hand. As will be seen, this does not appear to be the case for factor analytic studies of numerical facility.

II. Identifying a numerical factor

By definition, numerical facility involves an ability to work with numbers. However, there are numerous tasks that, in one form or another, involve the manipulation of numbers. All entail recognition of numerical symbols, but this is where much of the similarity ends. Some tests involving numbers are almost pure memory tasks (e.g., digit span on the Wechsler Adult Intelligence Scale; Wechsler, 1958). Others require manipulating numbers using only the basic rules of arithmetic, while still others entail more involved numerical and symbolic manipulation (e.g., algebraic operations). Some items involving number manipulation may require complex reasoning in some form or another (e.g., number series).

The types of items described above can be differentiated in many ways. One of the more obvious of these involves viewing them as varying on a continuum from more to less "automatic." Multiplying two one-digit numbers may, for many people, be no more than an exercise in recalling the multiplication table. However, the solution of problems involving arithmetic or mathematical reasoning is usually not as easily accomplished. Many of the studies presented below make a distinction between problems involving simple arithmetic and those requiring arithmetic reasoning. Following are factor analytic studies of mental abilities relating to numerical facility.

Thurstone's Primary Mental Abilities (PMA): Thurstone (1938) administered 57 tests to 218 volunteers at the University of Chicago. Using factor analysis, Thurstone found that 13 factors sufficiently explained the intercorrelations among the various tests. One of these he named the Number Factor (N). (Table 1 presents a list of factor abbreviations used in this text.) Nine of the ten tests involving the manipulation of numbers loaded on this factor at .35 or higher; tests restricted to simple arithmetic calculations loaded highest on N (.62-.81). Thurstone writes that "the insistence of the numerical factor makes it almost certain that it represents a unique ability" (p. 83).

Tests not restricted to arithmetic calculation tended to load highly on reasoning and other factors. Thus, Number Series loaded .50 on the Induction Factor, and Arithmetical Reasoning, a test containing "word problems," loaded .58 on Thurstone's R, an ill-defined factor involving restricted tasks of various types.

As noted earlier, much about the factor analytic technique is subjective. Thurstone constructed his factors to achieve something similar to what he later termed "simple structure" (Thurstone, 1947). However, it is not at all clear what constitutes simple structure (Nunnally, 1967). Since Thurstone's monumental study, several other factorial solutions have been proposed for the same data set. Kaiser (1960) reviews four of these and concludes that the correlations between the different solutions are high. "More particularly," he writes, "correlations between number factors defined by the different rotational solutions generally are higher than other correlations [.90-.97]" (p. 155). In other words, Thurstone's conclusion appears to hold across different factor analytic solutions.

Table 1

Abbreviations for Factors Used in this Review

D - Deductive Reasoning
M - Mathematics
N - Number
P - Perceptual
S - Spatial
V - Verbal

The Army Air Forces (AAF) Classification Batteries (Guilford and Lacey, 1947): In 1942 and 1943 the Army Air Forces Aviation Psychology Program administered 15 test batteries to classified and unclassified personnel. Four of the batteries were designated as classification batteries; the remainder were designed to tap various hypothesized traits such as judgment,

carefulness, and integration. A factor analysis was performed on each of these batteries, and factors were derived so as to be uncorrelated with each other (orthogonal).

The four classification batteries were somewhat similar in that they shared many of the same tests. The number of factors extracted from these batteries ranged between 10 and 14. Many of the remaining batteries differed greatly from one another in both the number and content of the tests that they contained. The number of factors extracted from the remaining eleven test batteries ranged between 6 and 13 with a mean of 9.

In any factor analytic study, the factors that are eventually extracted depend, in great measure, on the tests employed. Thus, for example, a motor learning factor is not expected to emerge from a test battery restricted to purely intellectual tasks. Nevertheless, despite the many differences among the batteries administered by the AAF, several results relating to numerical facility appeared consistently across all or most of the batteries. N emerged in factor analyses of three of the four classification batteries and in seven of the remaining eleven batteries.

Before we can proceed in greater detail, a brief description of the pertinent tests is warranted:

1. Mathematics A--An achievement test consisting of problems in advanced arithmetic, algebra, and trigonometry. It was expected that high school graduates would be able to answer most of these items correctly.

2. Mathematics B--A test that included one or two of the following:

   (a) Arithmetic Reasoning (1)--Items that attained some degree of difficulty but that required only minimal formal training in mathematics.

   (b) Arithmetic Reasoning (2)--Problems essentially similar to those in (1) but requiring more formal training in mathematics.

   (c) Numerical Approximation--Numerical operations involving many digits. Subjects were required to estimate the answer and to choose it from five available alternatives.

3. Numerical Operations Front--Simple subtraction and division problems. Subjects were required to indicate whether the solution presented was correct.

4. Numerical Operations Back--Simple subtraction and division problems. For each problem, subjects were required to choose the correct solution from five alternatives.

5. Dial and Table Reading--Subjects were presented with a test booklet that contained two tests; the two tests were considered separately. In the Dial Reading Test, subjects were shown various dials and were asked to read and report the numbers on those dials. In the Table Reading Test, subjects were asked to find appropriate quantities within tables.

The Numerical Operations tests (Front and Back) turned out to be the primary defining tests of N; they consistently showed their highest loadings on that factor, and they appeared in all but three of the ten batteries in which N emerged. Furthermore, in one of the three batteries in which N emerged and in which Numerical Operations were not given, the designation of N was questionable. Two of the tests loading highest on N in this battery did not contain problems involving numbers, nor did they include problems that otherwise appeared related to numerical facility.

The amount of variance in Numerical Operations scores explained by N ranged between .40 and .69 with a mean of .58. In only one case did Numerical Operations load significantly on a factor other than N. This was in the September 1944 classification battery in which Numerical Operations Front was found to load .25 on the Spatial Factor (S). Because this is a relatively small loading, and because Numerical Operations Back loaded exclusively on N, this result can be treated as a mere curiosity. It should be noted that the two Numerical Operations tests displayed similar loadings on N throughout the AAF batteries.

The patterns of loading for the other tests that are typically associated with N (Mathematics A and B, and Dial and Table Reading) are somewhat more complex. There generally did not appear to be differing patterns of loadings across these tests. However, the previous statement should be taken with caution. The data on which it is based is limited; most batteries contained only one of these mathematics tests. Listed below are some of the factor loading patterns exhibited by the mathematics tests:

1. Wherever N emerged and Mathematics A and/or B was given, the test loaded significantly on that factor. The amount of variance in all the forms of the mathematics tests explained by N ranged between .14 and .42 with a mean of .23. The two tests did not generally differ with respect to their loading on N.

2. Whenever a Verbal Factor (V) emerged and Mathematics A and/or B was given, the test loaded significantly on that factor. The amount of variance in the mathematics tests accounted for by V ranged between .07 and .34 with a mean of .14. However, where V was present and N was not, loadings on V tended to be higher. In other words, it appears that at least some of the variance in mathematics tests was shared by N and V.

3. In all but one battery, whenever Mathematics A and/or B was given it loaded significantly on a Deductive Reasoning Factor (D). Variance in mathematics tests explained by this factor ranged between .06 and .46 with a mean of .24. In other words, N and Deductive Reasoning factors explained similar amounts of variance in arithmetic reasoning and mathematics.

4. Dial and/or Table Reading tests appeared in four batteries. In every case, these tests loaded highest on N. The amount of variance accounted for by N in these tests was at or about .25. In all four batteries, these tests also loaded significantly on a Perceptual Factor (P) and S. Variance accounted for by S in these tests ranged between .12 and .23 and averaged .17. The variance in these tests explained by P ranged between .07 and .09. Dial and Table Reading appear to have similar magnitudes and patterns of loading across the four batteries.

5. In the only battery in which Number Series appeared, it loaded .47 on N and .36 on Deductive Reasoning.

Conclusions from the AAF batteries: A rough summary of the analyses of 15 AAF test batteries has been presented. As already noted, a great range of tests was represented in these batteries. Appropriately, the number and content of factors that emerged from these batteries often differed greatly. Despite these differences, a fairly clear picture emerged for tests involving the manipulation of numbers.

The highest loadings on N came from tests with items requiring only simple arithmetic computation. In the absence of these tests, N did not usually materialize. Almost 60% of the variance in the computational items could be explained by N.

Arithmetic reasoning and mathematics items correlated almost equally with N and Deductive Reasoning. These tests also displayed moderate correlations with V. This may indicate that the reasoning required in word problems is similar to that which is required in other areas as well. In other words, arithmetic reasoning, as opposed to arithmetic computation, may not be a discrete ability. This possibility is discussed in the following section. The fact that arithmetic computation did not significantly correlate with Deductive Reasoning or with any other reasoning factor suggests that it is a discrete ability.

Dial and Table Reading were associated with N more than they were with any other factor. Both of these tests were consistently related to P and S, as well.

Other factor analytic studies

Comrey (1949) gave 815 West Point graduates 13 tests of ability taken from the AAF batteries. He also obtained the

cadets' grades for several military and nonmilitary courses. As in previous studies, numerical operations (subtraction and division in Comrey's study) loaded exclusively on N (.52). Mathematics B also loaded on N (.52), but loaded almost equally on a Deductive Reasoning Factor. More moderate loadings were observed on V and on a Mechanical Reasoning Factor. Surprisingly, grades in mathematics did not load significantly on N. They did load on D (.63) and on a more general Grade Factor (.52), which is to be expected, as grades in mathematics are a subset of all grades. All the grades obtained for this study (history, English, physics, topography/graphs, tactics, mathematics) loaded significantly on D. Of these, only English loaded significantly on N (.26).

Chein (1939) gave 186 college students numerous tests of verbal, spatial, and numerical abilities. Almost all of these tests loaded significantly on V, and all but four loaded on a number factor. However, both verbal and and spatial tests loaded significantly on this factor. He gave no purely computational tests, and it is probably for this reason that a discrete N did not emerge. The number factor that did emerge in Chein's study resembles a general reasoning factor more than it does the N that appears in other studies.

Dudek (1948) also found a number factor, despite the absence of computational tests. In his study, Mathematics A and B loaded on N (.47 and .55, respectively) as did Dial and Table Reading. The population he used was somewhat less qualified (by flight school standards) than that used in the AAF batteries. When the same battery was given to 195 women, N did not emerge as clearly.

While Dudek's (1948) results are atypical, he was not alone in finding a number factor without the inclusion of tests that are purely computational (e.g., AAF Perceptual Battery I, AAF Perceptual Battery II). Nevertheless, most of the studies that include mathematical tests but do not include purely computational tests fail to find a clear N (e.g., AAF Classification Battery, Nov. 1943; AAF Memory Test I; AAF Memory Test II; AAF Mechanical Test; Chein, 1939; Garret, 1938). Thus, it appears that the appearance of a number factor that is separate from a reasoning factor is dependent on the inclusion of computational tests.

In all but one of the studies cited to this point (the AAF Classification Battery, Sept. 1944), tests involving simple computation loaded exclusively on N. In another study (Goodman, 1943), addition and multiplication were found to have moderate correlations with V (.37 and .30, respectively). This result is unique and may be attributable to the sample of engineers used in the study.

A confirmatory factor analytic study

Very few studies in the area of numerical facility have taken the confirmatory route. In one of these, Coombs (1941) hypothesized that arithmetic ability involves the capacity to manipulate symbols using prespecified rules. Coombs further hypothesized that arithmetic manipulation is a serial response and that it is related to perceptual speed. Consequently, he reasoned, the manipulation of any set of symbols should correlate highly with arithmetic ability.

Subjects in Coombs's study were administered items requiring manipulation of letters using rules that were consistent with his conceptualization of numerical ability. They were also given tests measuring seven of Thurstone's (1938) Primary Mental Abilities; among the abilities included were two types of reasoning (inductive and deductive), Verbal, Perceptual Speed, and Numerical. In this study, only tests of simple arithmetic were found to load primarily on N. Nevertheless, the correlations between tests of letter manipulation and those involving arithmetic operations ranged between .27 and .52 (N=223). Of 30 such correlations, 24 were .35 or higher. The correlation between the two tests best measuring P and those measuring N ranged between .13 and .28, with six of the ten correlations .15 or lower. Coombs also found an increase in correlation between tests of numerical ability and letter manipulation following practice on the latter. This makes sense when one considers that simple arithmetic is a highly practiced response for most individuals. In order to measure numerical facility, it is probably necessary to use items that have been highly practiced.

Coombs (1941) concluded that "the number factor is most clearly identified by very simple number tests," and that his results "are in agreement with the hypothesis that number ability is characterized by a facility in manipulating a symbolic system according to a specified set of rules" (pp. 188-189). He also concluded that the results do not support the hypothesis that numerical ability is a serial response. Vernon (1961) disputed Coombs's conclusions. He feels that the relatively small correlations between N and tasks not involving numbers indicate that number facility is number specific.

A common problem with studies such as the one noted above involves the distinction between aptitude and achievement. Vernon (1961) concluded that since the highest loadings on N involved actual arithmetic tasks, it is these tasks that are best suited to measure arithmetic ability. However, if achievement and aptitude are only moderately correlated, we would not necessarily expect aptitude measures of numerical ability (such as letter manipulation) to correlate primarily with achievement-type measures of arithmetic. In any event, the ideas

presented in Coombs's (1941) paper are important because they may provide a method to separate ability from achievement in the area of numerical facility.

III. Mathematical Ability

In this paper, the term 'mathematics problems' is used as a generic term for problems involving more than just rote arithmetic. This includes such problems as number series, "word problems," and the like. Generally, these problems display similar patterns of loading across various ability factors.

Forsyth and Ansly (1982) attempted to experimentally control for computational skills in the solution of mathematics problems. They administered mathematical reasoning problems to 567 ninth- and eleventh-grade high school students. In one condition, the students were to complete all the necessary computation by hand; in the other, they were to use calculators. No significant differences were found between the two groups. This was taken to indicate the minor role computational skills play in the solution of mathematical problems.

In many of the studies cited to this point, mathematical ability correlated with N as well as with other factors. This is generally seen as a result of the overlap between problems involving mathematical skill and those restricted to arithmetic computation. However, Forsyth and Ansly's (1982) result suggests that differences in computational skills may have little to do with differences in the ability to solve mathematical problems.

Werdelin and Stjernberg (1971) gave subjects arithmetic problems, number series, and nonmathematical logical reasoning. They found that the more difficult a problem was, the higher it tended to load on a general reasoning factor. In the Kit of Factor-Referenced Cognitive Tests (Ekstrom et al., 1976), the General Reasoning Factor is defined exclusively by tests involving the manipulation of numbers. However, none of these tests are purely computational.

Wrigley (1958) asked if there is a group factor above and beyond g (the general factor) in mathematical ability. He reviewed several contradictory studies and concluded that much of the confusion arose because inappropriate items were included in some mathematics tests. More specifically, many studies included geometry problems that load significantly on a spatial factor. In his study, three mathematics tests, algebra, geometry, and "problem arithmetic," were given in conjunction with spatial and computational tests. Both mathematics and arithmetic group factors emerged. However, while the computational tests loaded highly on N (.64-.75), the mathematics tests displayed only moderate correlations (.20-.36) with the Mathematics Factor (M). Of the mathematics tests, geometry showed the lowest correlation with M. Wrigley concludes that arithmetic and mathematics are differentiated and that once g is taken out, there is little left

for M to pick up. Nevertheless, he believes that there is some unique variance in M that is not associated with N or g.

Furneaux and Reese (1978) agree with Wrigley that there is an M factor that is separate from g. Their data suggests that the unique variance associated with M does not exist in every mathematics problem. In their study, they statistically differentiated problems that contained this unique variance from those that did not. However, they could not identify a conceptual basis for this differentiation.

IV. Automatic processes and numerical facility

The initial learning of arithmetic involves reasoning (Werdelin and Stjernberg, 1969). After a time, numerical operations become highly practiced to the point that for most educated adults they are "automatized." Consequently, it is possible that some of the individual differences in the performance of these operations may be a result of differences in ability to automatize.

Keats (1965) administered arithmetic, arithmetic reasoning, verbal, and perceptual tests to 160 college freshmen. He found a possible "automatic process" factor on which multiplication loaded .88. Because of the high loading, it is suspected that his automatic process factor is a variant of N.

As noted earlier, Coombs (1941) administered nonnumerical tests that involved manipulation of symbols according to a set of specified rules. The correlations between these tests and numerical tests were moderate. When subjects were given practice sessions on the nonnumerical tests, these correlations increased. In other words, nonnumerical symbolic manipulation shares more variance with numerical manipulation after practice than it does before practice. This would indicate that at least some of the variance in N can be explained by the ability to utilize a practiced response.

In a factorial investigation of perceptual speed, Bechtold (1947) administered perceptual, verbal, and numerical tests to 167 college students. The factor solution was oblique, with the first order factor N correlating highest (.42) with a first order factor defined as Speed of Recognition of Predetermined Symbols in the Contexts of Discrete Distractors. N correlated .74 with the second order factor of Perceptual Speed. In another study of perceptual speed, Werdelin and Stjernberg (1969) found that practice increased the correlation of tests of perception with N. They hypothesize that at least some of the variance explained by the perceptual speed factor is due to the test materials used.

V. Summary and Conclusions

Thurstone (1938) is confident that the ability to manipulate numbers is a "unique ability" (p. 83). He also feels there is some evidence that this ability is genetic in nature. French (1951) writes: "The number factor is the clearest of them all" (p. 115). He suggests that any test involving the manipulation of numbers will load highly on this factor. Carroll (1983) identifies N as appearing across numerous studies under the same and different labels.

It is clear that there is an ability associated with, and unique to, the manipulation of numbers. In most of the studies surveyed, the primary defining items of this facility were tests restricted to simple numerical manipulations. Tests consisting of simple arithmetic manipulations appear to correlate higher with N than do tests that are more complex (e.g., addition and subtraction tend to correlate with N to a greater degree than do multiplication and division).

Tests involving arithmetic reasoning consistently show loadings on N. These loadings tend to be moderate (.30 -.50) and are typically lower than these test's correlations with reasoning factors. Intuitively, one would think that it is the numerical manipulation requirement in these tests that brings about these correlations. However, there are indications that this may not be the complete explanation. First, the Forsyth and Ansley (1982) study of arithmetic reasoning found no difference in test scores between those who used calculators and those who did not. Second, some tests involving numbers and requiring no arithmetic manipulation (e.g., Dial Reading and Letter Manipulation) also tend to load moderately on N.

Some investigators (Furneaux and Reese, 1978; Wrigley, 1958) believe that there is a mathematical group factor separate from N and reasoning factors. The evidence for this is far from unequivocal. In any event, the amount of variance in mathematical ability that this factor may account for is probably very small.

Arithmetic operations have also been associated with automatic responses (i.e., the ability to respond quickly and appropriately in a stereotyped fashion). In several studies, such operations have correlated with perceptual speed. However, the speeded aspect of arithmetic operations is probably not unique to N and can be picked up by perceptual tests such as visual scanning.

Finally, there is the question of aptitude versus achievement. Practice usually enhances performance, and numerical operations and reasoning are no exceptions. The question is whether there would still be variance on these tests

were "life-practice-effects" controlled for. The answer is, "most probably yes." Thurstone (1938) points out that there have been individuals who very early on showed an exceptional capacity to manipulate numbers, and yet were no more than average with regard to other mental abilities. He and others (see Carroll, 1983, for references) have suggested that a significant amount of the variance in this ability may be due to genetic factors.

It may be possible to separate ability from achievement in numerical ability. Coombs (1941) administered tasks that had never been practiced and that, he thought, were similar to numerical operations. With greater practice, the correlations between these items and pure arithmetic items increased. It is possible that tests of arithmetic manipulation are, in part, tests of an individual's ability to benefit from practice over time.

Recommendations

Based on the review of the literature, it is recommended that three tests be used by the Johnson O'Connor Research Foundation to research the construct of numerical facility. The tests and the rationale for their use follows:

1. Simple arithmetic: Studies investigating numerical facility have found that N is best measured by tests of simple arithmetic manipulation. Consequently, any battery attempting to measure N must contain at least one such test. This test will be used both to measure N and as a marker for the other tests contained in the battery.

2. Simple arithmetic that involves some arithmetic reasoning: As arithmetic tests move from simple arithmetic manipulation toward arithmetic reasoning, they tend to load less on N and more on general reasoning factors. However, tests requiring moderate levels of arithmetic reasoning load moderately on N. It is possible that the simple arithmetic manipulation contained in these tests does not fully explain the variance in these tests associated with N. In order to comprehensively map the domain of N, it is probably necessary to employ tests that involve some reasoning as well as simple arithmetic.

3. Nonnumerical manipulation: Tests of simple arithmetic have the disadvantage of confounding ability with experience. In order to eliminate this confounding, it is necessary to construct items that require the use of rules that are similar to arithmetic but novel enough to eliminate the effects of practice.

References

Anastasi, A. (1976). Psychological testing (4th ed.). New York: Macmillan.

Bechtold, H. P. (1947). Factorial investigation of perceptual speed factor. American Psychologist, 2, 300-305.

Carroll, J. B. (1983). Studying individual differences in cognitive abilities: Through and beyond factor analysis. In R. F. Dillon and R. R. Schmeck (Eds.), Individual differences in cognition. New York: Academic Press.

Chein, I. (1939). An experimental study of verbal, numerical and spatial factors in mental organization. Psychological Record, 3, 71-94.

Comrey, A. L. (1949). A factorial study of achievement in West Point courses. Educational and Psychological Measurement, 9, 193-209.

Coombs, C. H. (1941). A factorial study of number ability. Psychometrika, 6, 161-189.

Dudek, F. J. (1948). The dependence of factorial composition of aptitude tests upon population differences among pilot trainees. Educational and Psychological Measurement, 8, 613-633.

Ekstrom, R. B., French, J. W., Harman, H. H., and Dermen, D. (1976). Manual for kit of factor-referenced cognitive tests. Princeton, N.J.: Educational Testing Service.

Forsyth, R. A., and Ansly, T. N. (1982). The importance of computational skills for answering items in a mathematical problem solving test: Implications for construct validity. Educational and Psychological Measurement, 42, 257-263.

French, J. W. (1951). The description of aptitude and achievement tests in terms of rotated factors. Psychometric Monographs Number 5. Chicago: University of Chicago Press.

Fruchter, B. (1948). The nature of verbal fluency. Educational and Psychological Measurement, 8, 33-47.

Furneaux, W. D., and Reese, R. (1978). The structure of mathematical ability. British Journal of Psychology, 69, 507-512.

Garret, H. E. (1938). Differential mental traits. Psychological Record, 2, 259-298.

Goodman, C. H. (1943). A factorial analysis of Thurstone's primary mental abilities test. Psychometrika, 8, 141-151.

Guilford, J. P., and Lacey, J. I. (Eds.). (1947). Printed Classification Tests. Army Air Forces Aviation Psychology Program Research Reports, No. 5. Washington, D.C.: United States Government Printing Office.

Kaiser, H. F. (1960). Varimax solution for primary mental abilities. Psychometrika, 25, 153-158.

Keats, J. A. (1965). An experimental study of cognitive factors. Australian Journal of Psychology, 17, 52-57.

Nunnally, J. (1967). Psychometric theory. New York: McGraw-Hill.

Thurstone, L. L. (1938). Primary mental abilities. Psychometric Monographs Number 1. Chicago: University of Chicago Press.

Thurstone, L. L. (1947). Multiple factor analysis. Chicago: University of Chicago Press.

Vernon, P. E. (1961). The structure of human abilities. London: Methuen.

Wechsler, D. (1958). The measurement and appraisal of adult intelligence (4th ed.). Baltimore: Williams and Wilkins.

Werdelin, I., and Stjernberg, G. (1969). On the nature of the perceptual speed factor. Scandinavian Journal of Psychology, 10, 185-192.

Werdelin I., and Stjernberg (1971). The relationship between and factor loadings of some visual-perceptual tests. Scandinavian Journal of Psychology, 12, 21-28.

Wrigley, J. (1958). The factorial nature of ability in elementary mathematics. British Journal of Educational Psychology, 28, 61-78.

JOHNSON & JOHNSON RESEARCH FOUNDATION
HUMAN ENGINEERING LABORATORY